

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-110840  
 (43)Date of publication of application : 30.04.1996

---

(51)Int.Cl. G06F 3/06  
 G06F 3/06

---

(21)Application number : 06-246107 (71)Applicant : NEC TELECOM SYST LTD  
 NEC CORP  
 (22)Date of filing : 12.10.1994 (72)Inventor : TAWARA HIROSHI  
 OZAKI YUICHI

---

(54) RESTORATION SYSTEM FOR MAGNETIC DISK DEVICE OF DOUBLE REDUNDANCY CONSTITUTION

(57)Abstract:

PURPOSE: To make the necessary time for copy in a needed disk device shorter than before by referring to a file update management table when a magnetic disk device of double redundancy constitution is put back to double redundancy operation from single independent operation.

CONSTITUTION: At the time of restoration from the single independent operation state of only a system 10A to the double operation state wherein a system 10B serves as a stand-by system only updated files are copied from a magnetic disk device 3A to a magnetic disk device 3B while referring to the file update management table 4 as a list of the names of files updated during the single independent operation on the side of the system 10A thereby synchronizing the pieces of information on the magnetic disk devices 3A and 3B with each other.

---

## CLAIMS

---

[Claim(s)]

[Claim 1] In a restoration system of a double-sized redundant configuration magnetic disk drive with a processing process which writes the same contents also in a magnetic disk drive of other systems while each two general purpose processor which carried out the redundant configuration carries out contents writing of a magnetic disk drive of a self-system A file name of a file in said magnetic disk of a self-system updated by shutdown of other systems during single-sided operation of only a self-system is accumulated in a file update

management tableBy referring to said file name accumulated in the file update management table at the time of operation restoration of other systemsA restoration system of a double-sized redundant configuration magnetic disk drive including a process which transmits file information corresponding to said file name to other systems from a self-systemand is copied into said magnetic disk of other systems.

[Claim 2]A restoration system of the double-sized redundant configuration magnetic disk drive according to claim 1 with which said file update table is set up in said magnetic disk of a self-system.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Industrial Application]About the restoration system of a double-sized redundant configuration magnetic disk driveespecially this invention is a telecommunications system etc. which need the high-reliability which is not allowed service interruptionand relates to the restoration system of the double-sized redundant configuration magnetic disk drive used when shifting to the restoration from the generating time of the piece system failure of a double-sized system.

[0002]

[Description of the Prior Art]The conventional double-sized redundant configuration magnetic disk drive has connected the systems 11A and 11B constituted by providing the magnetic disk drives 3A and 3B which are the general purpose processors 1A and 1Bthe main memory 2A and 2Band an auxiliary storage unitrespectively with the synchronous bus 5as shown in drawing 2. When the general purpose processor 1A of the system 11A is operating by this double-sized redundant configuration systemthe general purpose processor 1B of the system 11B is a waiting state. In that casethe magnetic disk drive 3A in the system 11A and the same entire contents are maintained by the redundant writing from the general purpose processor 1A of the system 11Aandas for the magnetic disk drive 3B in the system 11Bthe latest information is always stored. When the general purpose processor 1B of the system 11B is operatingThe general purpose processor 1A of the system 11A is a waiting statethe magnetic disk drive 3B and identical content in the system 11B are maintained by the redundant writing from the general purpose processor 1Bandas for the magnetic disk drive 3A in the system 11Athe latest information is always stored.

[0003]For examplewhen it stops by failure or a maintenance service in the middle of operation of the system 11Athe connection between the system 11A and the system 11B will be in a separation state. In the meantime the system 11B continues service employment by individual operation. Under the individual operation of such a single-sided systema difference arises by the contents of the information on the magnetic disk drive 3A in the system 11Aand the information

on the magnetic disk drive 3B in the system 11B. The moment that the shift to double-sized operation from 1-fold-sized operation is attained the redundant copy processing for abolishing the difference in such information content starts. In this redundant copy processing the volume copy from the magnetic disk drive 3B to the magnetic disk drive 3A of the system unit 11A by the side of the system 11B is performed. After the end of redundant copy processing the magnetic disk drives 3A and 3B of the systems 11A and 11B will have the same information mutually.

[0004]

[Problem to be solved by the invention] Since the volume copy between the magnetic disk drive 3A and 3B is performed in the redundant copy processing in the former mentioned above since there is no file update during 1-fold-sized (independent) operation and copy processing is performed also about the file which does not need to be copied there is a problem of wasting processing time for a copy excessive at the time of the restoration to double-sized operation.

[0005]

[Means for solving problem] In the restoration system of a double-sized redundant configuration magnetic disk drive with the processing process which writes the same contents also in the magnetic disk drive of other systems while as for the restoration system of this invention each two general purpose processor which carried out the redundant configuration carries out the contents writing of the magnetic disk drive of a self-system. The file name of the file in said magnetic disk of the self-system updated by the shutdown of other systems during single-sided operation of only a self-system is accumulated in the file update management table. By referring to said file name accumulated in the file update management table at the time of operation restoration of other systems the process which transmits the file information corresponding to said file name to other systems from a self-system and is copied into said magnetic disk of other systems is included.

[0006]

[Function] By the conventional restoration system volume copies are standard specifications and especially in the case of a mass disk a copy takes great time to them. Since only the file which had updating during 1-fold-sized operation carries out the redundant copy for a synchronization when it shifts to double-sized operation from 1-fold-sized operation the operating duty which a file takes can be managed with this invention in a necessary minimum short time.

[0007]

[Working example] Next this invention is explained in detail with reference to Drawings.

[0008] Drawing 1 is a block diagram showing one working example of this invention. In this example the file update management table 4 is formed in the magnetic disk drive 3A. Although the file update management table 4 does not exist on the magnetic disk drive 4A at the time of the usual double-sized operation when it changes into 1-fold-sized individual-operation state by failure or maintenance it creates on the magnetic disk drive by the side of operation (this example magnetic

disk drive 3A). And the name of the file is written in the file update management table 4 about the file which had updating in the file content during 1-fold-ized individual operation. The redundant copy of only a required part is performed referring to the file update management table 4 from 1-fold-ized individual operation at the moment of becoming possible to the magnetic disk drive (3B) of the side restored from the magnetic disk drive (3A) by the side of operation to double-ized operation.

[0009] Drawing 3 is a sequence chart which illustrates the programing operation from the moment of shifting to 1-fold-ized individual-operation state from double-ized operational status to the moment of restoring to a double-ized operation system again in this example. The moment the hard failure of one side occurred during double-ized operation the system which hard failure generated is separated automatically and 1-fold-ized (independent) operation is started. The parent process of a program gets to know that the point to point system carried out the failure occurrence by a system call (treating number 1) and tells a child process about that (treating number 2). The child process which received the notice from the parent process creates the file update management table 4 (treating number 3). and in 1-fold-ized operation -- a file name -- "name "a" of the file in which updating generated the parent process to the child process when the renewal of the file "a" occurred" is transmitted (treating number 4). A child process writes file name "a" which received transmission in the file update management table 4 (treating number 5). If the file "b" is updated similarly file name "b" will be transmitted to a child process (treating number 6) and the received child process will be added to the file update management table 4 (treating number 7).

Restoration \*\*\*\* gets across to a parent process by a system call at the moment of hard failure having been restored after that and changing into the state which can be double-ized operated (treating number 8). According to this a parent process carries out message transmission (treating number 9) of the opposite side system being in the state which can be restored to a child process. In order that a child process may synchronize the file content of a double-ized magnetic disk drive according to this message a file copy is carried out to the magnetic disk drive (3B) by the side of restoration from the magnetic disk drive (3A) by the side of operation (treating number 10) and a response is returned to a parent process after the completion of a copy (treating number 11). Since the file copied at this time is only a part of the file name in the file update management table 4 the time required of the synchronization copy for restoration can be managed with necessary minimum and can suppress unnecessary increase of the load of a general purpose processor.

[0010]

[Effect of the Invention] As explained above in this invention the name of each file updated during 1-fold-ized operation is stored in the file update management table and a synchronization copy is performed at the moment of a system being restored in the state which can be double-ized operated referring to this file update management table.

Therefore the copy time required in the transition state to double-izing can be suppressed less than before.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is a block diagram of one working example of this invention.

[Drawing 2] It is a block diagram which illustrates a conventional system.

[Drawing 3] It is a sequence chart which illustrates operation of the redundant copy processing in working example of this invention.

[Explanations of letters or numerals]

1A and 1B General purpose processor

2A and 2B Main memory

3A3B magnetic disk drive

4 File update management table

5 Synchronous bus

10A10B11A and 11B System

---

(19) 日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11) 特許出願公開番号

特開平8-110840

(43) 公開日 平成 8 年 (1996) 4 月 30 日

|                           |         |        |     |        |
|---------------------------|---------|--------|-----|--------|
| (51) Int.Cl. <sup>6</sup> | 識別記号    | 庁内整理番号 | F I | 技術表示箇所 |
| G 0 6 F 3/06              | 3 0 4 E |        |     |        |
|                           | 3 0 6 D |        |     |        |

審査請求 有 請求項の数 2 O L (全 4 頁)

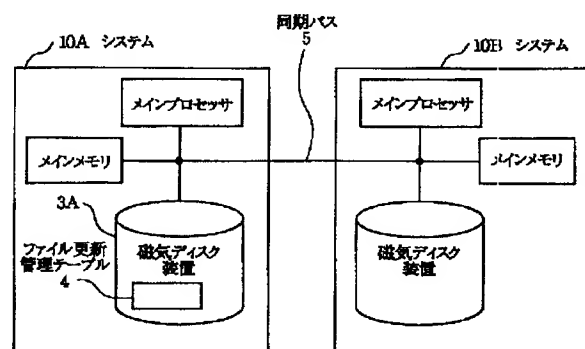
|           |                         |          |  |
|-----------|-------------------------|----------|--|
| (21) 出願番号 | 特願平6-246107             | (71) 出願人 | 000232106<br>日本電気テレコムシステム株式会社<br>神奈川県川崎市中原区小杉町 1 丁目 403 番地 |
| (22) 出願日  | 平成 6 年 (1994) 10 月 12 日 | (71) 出願人 | 000004237<br>日本電気株式会社<br>東京都港区芝五丁目 7 番 1 号                 |
|           |                         | (72) 発明者 | 田原 弘志<br>東京都港区芝五丁目 7 番 1 号 日本電気株式会社内                       |
|           |                         | (72) 発明者 | 尾崎 雄一<br>神奈川県川崎市中原区小杉町一丁目 403 番地 日本電気テレコムシステム株式会社内         |
|           |                         | (74) 代理人 | 弁理士 京本 直樹 (外 2 名)  |

(54) 【発明の名称】 2 重化冗長構成磁気ディスク装置の復旧方式

(57) 【要約】

【目的】 2 重化冗長構成磁気ディスク装置において、1 重化単独運転から 2 重化冗長運転に復旧するとき、ディスク装置のコピーが必要になる。このコピーの所要時間を、ファイル更新管理テーブルを参照することにより、従来よりも短縮させる。

【構成】 システム 10 A のみの 1 重化単独運転状態からシステム 10 B を待機系とした 2 重化運転状態への復旧時、磁気ディスク装置 3 A と磁気ディスク装置 3 B とを同期化するため、システム 10 A 側で 1 重化単独運転中にファイル更新のあったファイル名の一覧表であるファイル更新管理テーブル 4 を参照しながら更新のあったファイルのみ磁気ディスク装置 3 A から磁気ディスク装置 3 B へコピーし、磁気ディスク装置 3 A および 3 B 内の情報を同期化する。



**【特許請求の範囲】**

【請求項1】 冗長構成した2系統の各メインプロセッサが自系統の磁気ディスク装置の内容書き込みを実施するとともに同じ内容を他系統の磁気ディスク装置にも書き込む処理過程をもつ2重化冗長構成磁気ディスク装置の復旧方式において、他系統の運転中止により自系統のみの片側運転中に更新された自系統の前記磁気ディスク中のファイルのファイル名をファイル更新管理テーブルに蓄積しておき、他系統の運転復旧時にそのファイル更新管理テーブルに蓄積されている前記ファイル名を参照する事により、自系統から前記ファイル名に対応するファイル情報を他系統に転送して他系統の前記磁気ディスク中へコピーする過程を含むことを特徴とする2重化冗長構成磁気ディスク装置の復旧方式。

【請求項2】 前記ファイル更新テーブルは、自系統の前記磁気ディスク内に設定される請求項1記載の2重化冗長構成磁気ディスク装置の復旧方式。

**【発明の詳細な説明】****【0001】**

【産業上の利用分野】本発明は2重化冗長構成磁気ディスク装置の復旧方式に関し、特にサービス中断が許されない高信頼性を必要とする情報通信システムなどで、2重化システムの片系統故障の発生時からその復旧に移行する際に使用される2重化冗長構成磁気ディスク装置の復旧方式に関する。

**【0002】**

【従来の技術】従来の2重化冗長構成磁気ディスク装置は、図2に示すように、それぞれメインプロセッサ1Aおよび1B、メインメモリ2Aおよび2B、補助記憶装置である磁気ディスク装置3Aおよび3Bを具備して構成されたシステム11Aおよび11Bを同期バス5で接続してある。この2重化冗長構成システムで、システム11Aのメインプロセッサ1Aが運転している場合には、システム11Bのメインプロセッサ1Bは待機状態である。その場合、システム11B内の磁気ディスク装置3Bは、システム11Aのメインプロセッサ1Aからの冗長書き込みにより、システム11A内の磁気ディスク装置3Aと全くの同じ内容が維持され、いつも最新情報が格納されている。またシステム11Bのメインプロセッサ1Bが運転している場合には、システム11Aのメインプロセッサ1Aは待機状態で、システム11A内の磁気ディスク装置3Aはメインプロセッサ1Bからの冗長書き込みにより、システム11B内の磁気ディスク装置3Bと同一内容が維持され、いつも最新情報が格納されている。

【0003】たとえばシステム11Aの運転途中に故障または保守作業により停止した場合、システム11Aとシステム11Bとの接続は切り離し状態となる。その間、システム11Bが単独運転でサービス運用を続ける。このような片側システムの単独運転下では、シス

テム11A内の磁気ディスク装置3Aの情報とシステム11B内の磁気ディスク装置3Bの情報との内容に差異が生じる。1重化運転から2重化運転への移行が可能となった瞬間から、そのような情報内容の差異をなくするための冗長コピー処理がスタートする。この冗長コピー処理では、システム11B側の磁気ディスク装置3Bからシステム装置11Aの磁気ディスク装置3Aへのボリュームコピーが実行される。冗長コピー処理の終了後には、システム11Aおよび11Bの磁気ディスク装置3Aおよび3Bは互いに同一の情報を持つことになる。

**【0004】**

【発明が解決しようとする課題】上述した従来における冗長コピー処理では、磁気ディスク装置3Aおよび3B間でのボリュームコピーを行っているため、1重化（単独）運転中にファイル更新が無く、コピーする必要のないファイルについてもコピー処理が行われるので、2重化運転への復旧時に余分なコピーのために処理時間を浪費してしまうという問題点がある。

**【0005】**

【課題を解決するための手段】本発明の復旧システムは、冗長構成した2系統の各メインプロセッサが自系統の磁気ディスク装置の内容書き込みを実施するとともに同じ内容を他系統の磁気ディスク装置にも書き込む処理過程をもつ2重化冗長構成磁気ディスク装置の復旧方式において、他系統の運転中止により自系統のみの片側運転中に更新された自系統の前記磁気ディスク中のファイルのファイル名をファイル更新管理テーブルに蓄積しておき、他系統の運転復旧時にそのファイル更新管理テーブルに蓄積されている前記ファイル名を参照する事により、自系統から前記ファイル名に対応するファイル情報を他系統に転送して他系統の前記磁気ディスク中へコピーする過程を含むことを特徴とする。

**【0006】**

【作用】従来の復旧方式ではボリュームコピーが標準仕様であり、特に大容量ディスクの場合、コピーに多大な時間を要する。本発明では、1重化運転から2重化運転に移行した場合、1重化運転中に更新のあったファイルだけ同期のための冗長コピーを実施するので、ファイルに要する運転負荷が必要最低限の短時間で済む。

**【0007】**

【実施例】次に、本発明について図面を参照して詳細に説明する。

【0008】図1は本発明の一実施例を示すブロック図である。本実施例では、磁気ディスク装置3A内にファイル更新管理テーブル4を設けてある。通常の2重化運転時には、ファイル更新管理テーブル4は磁気ディスク装置4A上に存在しないが、故障やメンテナンスによる1重化単独運転状態になったときに、運転側の磁気ディスク装置上に作成する（本実施例では磁気ディスク装置3A）。そして、1重化単独運転中にファイル内容に更

新があったファイルについて、そのファイルの名前をファイル更新管理テーブル4に書き込んでいく。さらに、1重化単独運転から2重化運転に可能となった瞬間、運転側の磁気ディスク装置(3A)から復旧側の磁気ディスク装置(3B)に対し、ファイル更新管理テーブル4を参照しながら必要な分だけの冗長コピーを実行する。

【0009】図3は、本実施例において、2重化運転状態から1重化単独運転状態へ移行した瞬間から、再び2重化運転システムに復旧する瞬間までのプログラム動作を例示するシーケンスチャートである。2重化運転中に片側のハード故障が発生した瞬間、ハード故障が発生した系は自動的に切り離され、1重化(単独)運転が開始される。プログラムの親プロセスは、対向システムが故障発生したことをシステムコール(処理番号1)により知り、その旨を子プロセスに知らせる(処理番号2)。親プロセスから通知を受けた子プロセスは、ファイル更新管理テーブル4を作成(処理番号3)する。そして1重化運転中にファイル名"a"というファイルの更新が発生したら、親プロセスは子プロセスに対して更新が発生したファイルの名前"a"を送信する(処理番号4)。子プロセスは送信を受けたファイル名"a"をファイル更新管理テーブル4に書き込む(処理番号5)。同様に"b"というファイルが更新されたら、ファイル名"b"を子プロセスへ送信し(処理番号6)、受けた子プロセスはファイル更新管理テーブル4に書き加える(処理番号7)。その後ハード故障が修復され2重化運転可能状態になった瞬間、システムコールにより親プロセスに復旧可能が伝わる(処理番号8)。これに応じて親プロセスは子プロセスに対し、対向側システムが復旧可能状態であることをメッセージ送信(処理番号9)

する。子プロセスは、このメッセージに応じて2重化磁気ディスク装置のファイル内容を同期化するために、運転側の磁気ディスク装置(3A)から復旧側の磁気ディスク装置(3B)へファイルコピーを実施(処理番号10)し、コピー完了後、親プロセスへ応答を返す(処理番号11)。このときにコピーされるファイルは、ファイル更新管理テーブル4にあるファイル名の分だけであるから、復旧のための同期化コピーの所要時間は必要最少限で済み、メインプロセッサの負荷の不必要な増大を抑えることができる。

【0010】

【発明の効果】以上に説明したように本発明によれば、1重化運転中に更新された各ファイルの名称をファイル更新管理テーブルに格納しておき、システムが2重化運転可能状態に復旧した瞬間に、このファイル更新管理テーブルを参照しながら同期化コピーを行うことにより、2重化への遷移状態でのコピー所要時間を従来よりも少なく抑えることができる。

【図面の簡単な説明】

【図1】本発明の一実施例のブロック図である。

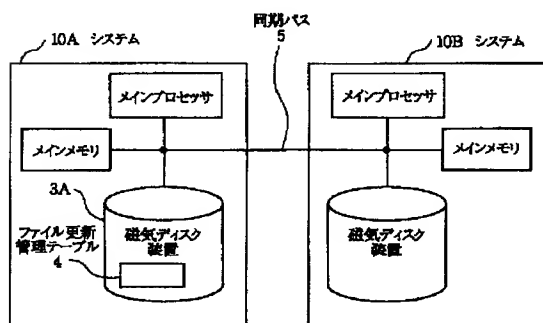
【図2】従来方式を例示するブロック図である。

【図3】本発明の実施例における冗長コピー処理の動作を例示するシーケンスチャートである。

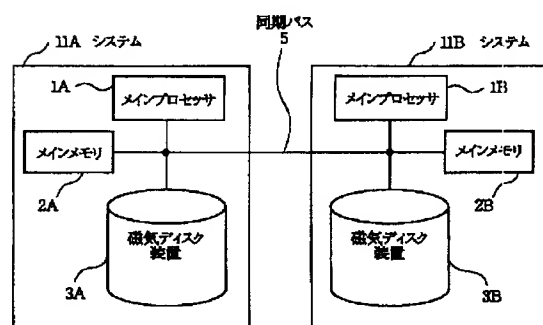
【符号の説明】

- 1A, 1B   メインプロセッサ
- 2A, 2B   メインメモリ
- 3A, 3B   磁気ディスク装置
- 4       ファイル更新管理テーブル
- 5       同期バス
- 10A, 10B, 11A, 11B   システム

【図1】



【図2】





【図3】

